Yet Another Proof of Cantor's Theorem

N. Raja

School of Technology & Computer Science
Tata Institute of Fundamental Research
Mumbai 400 005, INDIA
raja@tifr.res.in
http://www.tcs.tifr.res.in/~raja

Abstract: We present various proofs of Cantor's theorem in set theory: namely that the cardinality of the power set of a set X exceeds the cardinality of X, and in particular the continuum is uncountable. One of the proofs we present is inspired by Yablo's non-self-referential Liar's paradox, and it seems to bear a *dual* relationship to yet another proof.

Keywords: Cantor's theorem; Yablo's Paradox; Set theory; Proof Complexity

Mathematics Subject Classification: 03E99; 03F99

1 Introduction

Cantor's theorem – that for no set there is a function mapping its members onto all its subsets – is one of the most fundamental theorems in set theory and in the foundations of mathematics. It is interesting to explore different ways of proving such a basic theorem, possibly using a minimal or non-standard repertoire of basic constructs or reasoning mechanisms.

We begin by recalling Cantor's diagonalization proof, and note that it exhibits a subset which is *left-over* by any onto mapping from any set to its powerset. The traditional diagonalization proof constructs such a subset using the negation operator.

We introduce Yablo's non-self-referential Liar's paradox, and present a different proof of Cantor's theorem inspired by Yablo's paradox. This proof constructs another *left-over* subset which does not require invoking the negation operation for its definition.

We then discuss various aspects of the proof, and indicate similarities with two other paradoxes. We outline other proofs provided by Cantor; and finally show yet another proof which may in some sense be considered *dual* to the negation-free proof.

2 Cantor's Diagonalization Proof

We recall Cantor's diagonalization proof of his eponymous theorem.

Theorem 2.1 Cantor's Theorem: *For any set, there is no function mapping its members onto all its subsets.*

Proof [2, 3]: For any set X, let $P(X)$ denote the power set of X, i.e. $P(X) = \{T|T \subseteq X\}$. Suppose that the cardinality of X is equal to the cardinality of $P(X)$. This means that a one-to-one correspondence can be established between X and $P(X)$. For any $x \in X$ and $T_x \in P(X)$, let (x, T_x) denote pairs of elements established by the one-to-one correspondence. Now consider the set $D = \{y|y \notin T_y\}$. Clearly $D \in P(X)$, and D differs from every set T_y with respect to the element y. Thus any one-to-one correspondence omits the set D. Q.E.D.

Notice that the construction of the set D, which is *left-over* by any one-to-one mapping, involves the use of negation in stating $y \notin T_y$.

3 Paradox without Circularity

Yablo's paradox [13, 14, 15] is a non-self-referential Liar's paradox. Before the formulation of Yablo's paradox, all known paradoxes in logic seemed to require circularity in an unavoidable way. Each of them used either direct self-reference, or indirect loop-like self-reference. So, it appeared as though self-reference was a necessary condition for the construction of paradoxical sentences. Yablo's paradox demonstrated that this was not the case. We provide a brief outline of it in this section.

Consider the following infinite sequence of sentences S_i where the indices 'i, j, k' range over natural numbers:

$$(S_i) \;:\; \text{For all } \; j > i, \; S_j \; \textit{is untrue}$$

Note that, in the above sequence of statements, each statement quantifies only over statements which occur later in the sequence.

Suppose S_k is true for some k. Then S_{k+1} is false, and so are all subsequent statements. As all subsequent statements are false, S_{k+1} is true, which is a contradiction. So S_k is false for all k. Looking at any particular i, this in turn means that S_i in fact holds, which is a contradiction. Thus Yablo's paradox provides a sequence of statements such that none of them ever refer to themselves even in an indirect way, yet they are all both true and false.

4 Another Proof of Cantor's Theorem

Theorem 4.1 (Cantor's Theorem) *The cardinality of the power set of a set X exceeds the cardinality of X, and in particular the continuum is uncountable.*

Proof [9]: Let X be any set, and $P(X)$ denote the power set of X. Assume that it is possible to define a one-to-one mapping $M : X \leftrightarrow P(X)$

Define $s_0, s_1, s_2, ...$ to be a trace, where the first element of the trace is any arbitrary $s_0 \in X$, and all further elements s_j where $j > 0$, of the trace are such that $s_j \in M(s_{j-1})$

Define $t \in X$ to be a simple element, if all possible traces beginning with t terminate. Note that a trace $s_0, s_1, s_2, ..., s_f$ terminates at s_f if $M(s_f)$ is the empty set.

Define $N = \{t \in X | \text{t is a simple element}\}$

The set N, which is a subset of X, cannot lie in the range of M. Suppose there exists an $n \in X$ such that $M(n) = N$, then n should be a simple element since all traces beginning with element n also terminate. Thus $n \in N$, but then n is no longer a simple element, since not all traces beginning with n are terminating traces (e.g. "n, n, n, ..." is one such non-terminating trace).

Thus the set N is out of the range of mapping M. Q.E.D.

There is no explicit negation involved in the definition of the set N.

5 Generalizing Cantor's Agument

In this section we shall first see how the basic idea of Cantor's argument in the construction of the set D can be generalized in analogy with Yablo's non-self-referential liar's paradox. Such a generalization would give rise to the set N in the negation-free proof.

Extend the definition of a simple element to the notion of a k-simple element as follows.

Define $t \in X$ to be a k-simple element when for $k > 0$, there is no sequence $s_1, \dots, s_k$ such that $s_1 \in M(t), \dots, s_k \in M(s_{k-1})$ and $t \in M(s_k)$.

Let N_k be the set of all k-simple elements of X. Then an easy argument shows that there can be no $n \in X$ such that $N_k = M(n)$.

Cantor's argument has N_0, where t is 0-simple when $t \notin M(t)$. In the negation-free proof, the sets N_k are intermediaries, before leading up to the set N. We have replaced the circles mentioned in N_k by '*omega*', i.e. in analogy with Yablo's paradox, we have opened these circles, to construct N.

Does the set N use negation? It is perhaps not immediately obvious that the set N does not use negation. It is defined as:

$$N = \{t \in X | \ t \text{ is a simple element}\}$$

By definition, $t \in X$ is a simple element when all possible traces beginning with t terminate. Is the statement "all possible traces beginning with t terminate" negation-free? To settle this question we would have to rewrite it as a first-order formula. An obvious rewriting, which comes out of the definition of k-simple above, would correspond to: "there is **no** non-terminating sequence $s_1, s_2, \dots$ such that $s_1 \in M(t), \dots, s_k \in M(s_{k-1}), \dots$" Negation seems to occur explicitly in the above formula. But there is also another negation implicit in the notion of a **non**-terminating sequence above.

That the set N is negation-free can be shown as follows. We can also rewrite the statement "all possible traces beginning with s_0 terminate" as a first-order formula in another way. For $k \geq 0$, let an element s_0 in X be called simple when for each possible sequence (beginning with s_0): $s_0, s_1, \dots, s_k$ ($\forall i \ s_{i+1} \in M(s_i)$) there exists a j such that $M(s_j) = \phi$. This characterization is negation-free.

6 Hypergames and Grounded Classes

The reasoning involved in the negation-free proof also resembles the reasoning involved in establishing two well known paradoxes, viz., the *Hypergame* paradox and the *Mirimanoff's* paradox. The following sub-sections contain an outline of these paradoxes.

6.1 Hypergame Paradox

The Hypergame Paradox, also known as Zwicker's Paradox, was formulated by William Zwicker [5, 11, 17] in game theory.

Definition 6.1 (Two Player Finite Game): *A two-player game is defined to be finite if it satisfies the following conditions:*

1. *Two players, A and B, move alternately, A going first. Each has complete knowledge of the other's moves.*
2. *There is no chance involved.*
3. *There are no ties, i.e. when a play of the game is complete, there is one winner.*
4. *Every play ends after finitely many moves.*

Definition 6.2 (Hypergame): *The game Hypergame is a two-player game with the following rules:*

1. *On the first move, player A names any finite game F (called the subgame).*
2. *The players then proceed to play F, with B playing the role of A while F is being played.*
3. *The winner of the play of the subgame is declared to be the winner of the play of Hypergame.*

The Hypergame paradox is brought out by the question: Is Hypergame finite? As Hypergame satisfies the four conditions required for finite games, it is finite. If Hypergame is finite then player A can choose Hypergame as the finite game F of the first move. Now player B can name Hypergame as the first move. This process can lead to an infinite play, contrary to the assumption that Hypergame is finite.

6.2 Mirimanoff's Paradox

Mirimanoff's Paradox, also known as the Paradox of the Class of All Grounded Classes, was formulated by Dmitri Mirimanoff [6, 7, 8, 16], in set theory.

Definition 6.3 (Grounded Class): *A class X is said to be a grounded class when there is no infinite progression of classes X_1, X_2, ... (not necessarily all distinct) such that $\ldots \in X_2 \in X_1 \in X$.*

Definition 6.4 (Class of all Grounded Classes): *Let Y be the class of all grounded classes.*

Mirimanoff's Paradox is brought out by the question: Is Y, the class of all grounded classes, itself grounded? Let us assume that Y itself is a grounded class. Hence $Y \in Y$ and so we have $\ldots Y \in Y \in Y \in Y$ contrary to groundedness of Y. Therefore Y is not a grounded class. If on the other hand Y is not grounded, then there is an infinite progression of classes X_1, X_2, ... such that, $\ldots \in X_2 \in X_1 \in Y$. Since $X_1 \in Y$, X_1 is a grounded class. But then $\ldots \in X_2 \in X_1$, which means X_1 in turn is not grounded, which is impossible since $X_1 \in Y$.

7 Cantor's Other Proofs

In this section, we briefly sketch Cantor's two other proofs for the uncountability of the continuum [1, 2, 3, 10].

Theorem 7.1 (Uncountability of the Continuum) *There cannot be any one-to-one correspondence between the natural numbers and the real numbers.*

Cantor's Proof by Diagonalization [2, 3]: Consider the real numbers between zero and one, represented by infinite decimal expansions. Any attempt to construct a one-to-one correspondence between the the natural numbers and the reals will fail for the following reason. For any one-to-one correspondence we can construct a real number that is an infinite expansion which is different from every other real number in the range of the mapping. This can be done by making the number constructed differ from the first number of the mapping in the first decimal place; differ from the second number of the mapping in the second decimal place; and by continuing in this way to obtain an infinite decimal which is different from every other real number in the range of the mapping. Q.E.D.

Of course, one could say that the above diagonal argument is also a negation-free proof, because the process of swapping digits in the decimal expansion of a real number need not be thought of as negation.

Cantor's very first proof of the uncountability of the reals did not use diagonalization [1, 10]. It also does not depend on the fact that the real numbers can be represented by infinite decimal expansions. However it makes use of the topological properties that follow from the axiomatic characterization of the real numbers. This proof is not so widely known as the proof by diagonalization. We provide a brief sketch of it here.

Cantor's Proof without Diagonalization [1, 10]: Assume that an onto mapping from natural numbers to the reals gives the sequence $a_0, a_1, a_2, a_3, \ldots$ of real numbers. Let C_0 be a closed interval that does not contain a_0. Let C_1 be a closed subinterval of C_0 such that C_1 does not contain a_1. Continue this procedure to obtain an infinite nested sequence of closed intervals, $C_0 \supseteq C_1 \supseteq C_2 \supseteq \ldots$, that eventually excludes all the a_i's. Let r be a point that belongs to the intersection of all the C_i's. The real number r is different from all of the a_i. Q.E.D.

8 Yet Another Proof

We now exhibit yet another proof of the uncountability of the continuum. It is interesting to note that the following proof bears a kind of *dual* relationship to the negation-free proof. Recall that the negation-free proof uses the bijection $M : X \leftrightarrow P(X)$ in the direction $M : X \rightarrow P(X)$ in order to construct nondeterministic traces satisfying certain conditions, and relies on the empty set ($\phi \in P(X)$) in order to characterize their termination. On the other hand, the following proof uses the mapping M in the reverse direction, viz., $M : P(X) \rightarrow X$, and begins with the empty set in order to construct chains satisfying certain conditions. It may be possible to formalize these informal observations about the relationship between the two proofs in a rigorous way in a categorical framework [15].

We begin by stating an important property of posets, which will be required in the proof of the uncountability of the continuum.

Lemma 8.1 (Kurepa's Lemma) *If $P = (P, \leq_P)$ is a poset, and the tree $\sigma P = (\sigma P, \leq_{\sigma P})$ is defined as the set of ascending sequences of elements of P ordered by end-extension, then there is no order preserving, one-to-one mapping $f : \sigma P \to P$.*

Proof [4, 12]: Consult from Kurepa [4]. For a proof of a more general version of the above theorem, consult from Todorečević–Väänänen [12]. Q.E.D.

Theorem 8.2 (Uncountability of the Continuum) *There cannot be any one-to-one correspondence between the natural numbers and the real numbers.*

Proof: Let $X = \{a_0, a_1, \ldots\}$ be any set, and $P(X)$ denote the power set of X. Assume that it is possible to define a one-to-one mapping $M : P(X) \to X$.

Use M to construct chains in $P(X)$ and X such that:

$$\begin{aligned} a_0 &= M(\phi) \\ a_1 &= M(\{a_0\}) \\ a_2 &= M(\{a_0, a_1\}) \\ &\ldots \\ a_\omega &= M(\{a_0, a_1, \ldots\}) \\ a_{\omega+1} &= M(\{a_0, a_1 \ldots, a_\omega\}) \\ &\ldots \end{aligned}$$

By Kurepa's lemma, X has to be a proper class.

This contradicts the assumption that X is a set, hence a one-to-one mapping M cannot exist. Q.E.D.

9 Conclusion

Fundamental theorems are fascinating phenomena on their own right. Equally fascinating are paths which reconstruct their proofs using a minimal or even a non-standard repertoire of basic constructs and reasoning mechanisms. We have traversed such a path in this paper, and presented different proofs of a theorem which marked the beginnings of transfinite set theory. We have indicated how two of the proofs bear a special relationship to one another, and that it may be possible to to unify them in a common framework. We are exploring such a framework, and its further implications such as a possible mechanism which might automatically generate them from each other.

Acknowledgments

Thanks to Kosta Dosen, Zoran Petric, Jouku Väänänen and Noson Yanofsky for critically reading and commenting on preliminary versions of this paper. Thanks also to Margaret D'Souza for typing and proofchecking this paper.

References

[1] Cantor, G.: Über eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen, Journal für die reine und angewandte Mathematik, 77 (1874) 258–262.

[2] Cantor, G.: Über eine elementare Frage der Mannigfaltigkeitslehre, Jahresbreicht der Deutschen Mathematiker-Vereinigung, 1 (1891) 75–78.

[3] Crossley, J.N.: What is Mathematical Logic? Oxford University Press (1972) 4–5.

[4] Kurepa, D., Ensembles ordonnés et leur sous-ensembles bien ordonnés, C.R. Acad. Sci. Paris Ser. A 242 (1956), 116(1) (1995) 2202–2203.

[5] Miscellanea, The American Mathematical Monthly, 90 (1983) 390.

[6] Montague, R., On the Paradox of Grounded Classes, Journal of Symbolic Logic, 20 (1955) 140.

[7] Quine, W.V.O., Mathematical Logic, Revised Edition, Harvard University Press (1955) 130–131.

[8] Quine, W.V.O., Set Theory and its Logic, Harvard University Press (1963) 36–37.

[9] Raja, N., A Negation-Free Proof of Cantor's Theorem, Notre Dame Journal of Formal Logic, 46 (2005) 231–233.

[10] Rucker, R.: Infinity and the Mind: The Science and Philosophy of the Infinite, Birkhäuser (1982) 265–267.

[11] Smullyan, R., 5000 B.C. and Other Philosophical Fantasies, St. Martin's Press, New York (1983) 40–41.

[12] Todorčević, S., Väänänen, J., Trees and Ehrenfeucht–Fraissé games, Annals of Pure and Applied Logic 100 (1999) 69–97.

[13] Yablo, S.: Truth and Reflection, Journal of Philosophical Logic, 14 (1985) 297–349.

[14] Yablo, S.: Paradox without Self-reference, Analysis, 53 (1993) 251–252.

[15] Yanofsky, N.S.: A Universal Approach to Self-Referential Paradoxes, Incompleteness and Fixed Points, Bulletin of Symbolic Logic, 9 (2003) 362–386.

[16] Yuting, S., Paradox of the Class of All Grounded Classes, Journal of Symbolic Logic, 18 (1953) 114.

[17] Zwicker, W.S., Playing Games with Games: the Hypergame Paradox, American Mathematical Monthly, 94 (1987) 507–514.